DOCUMENT RESUME

ED 152 039 EC 110 029

AUTHOR	Krasner, William
TITLE	Labeling the Children.
INSTITUTION	National Inst. of Mental Health (DHEW), Rockville, Md.
REPORT NO	ADM 77-466
PUB DATE	77
GRANT	MH25044
NOTE	25p.
AVAILABLE FROM	Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 (Stock No. 017-024-00645-9)
EDRS PRICE	MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS	Conceptual Schemes; *Culture Free Tests; Elementary Secondary Education; *Evaluation Methods; *Labeling (of Persons); *Mentally Handicapped; *Minority Groups; *Models; Research Reviews (Publications); Student Evaluation; Student Placement; Test Bias

ABSTRACT

The report describes research on the effects of labeling children from minority groups as retarded and includes a review of a system of multiculturalistic pluralistic assessment (SOMPA), an instrument for evaluating the abilities and potentialities of children based on different aspects of performance. Listed among findings of the Riverside study, an 8-year study on the prevalence rates of mental retardation in a community, are that the public schools were the primary labelers, children labeled as mentally retarded by the schools were less deviant than those labeled by other agencies, and that the poor were consistently overlabeled. Explained are the three conceptual models for assessment incorporated in SOMPA: the medical model, the social system model, and the pluralistic model. Issues basic to the proper assessment, classification, and placement of children in the public schools are noted to include the requirement that testing take into account a child's language and cultural and racial origins and be nondiscriminatory. (SBH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

NATIONAL INSTITUTE OF MENTAL HEALTH

10,000

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED152039

EC110029

Labeling the children.

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service • Alcohol, Drug Abuse, and Mental Health Administration

Research Grant: MH 25044

Investigator: Jane R. Mercer, Ph.D.
Professor and Chair, Department of Sociology
University of California
Riverside, California 92502

Author: William Krasner

Date of Interview: August 1976

Mental Health Studies and Reports Branch
Division of Scientific and Public Information
National Institute of Mental Health
5600 Fishers Lane
Rockville, Maryland 20857

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

Stock No. 017-024-00648-9

☆ U.S. GOVERNMENT PRINTING OFFICE : 1977 O—243-928

001513

FOREWORD

Many children in our society are labeled according to the scores that they receive on various psychological tests and other diagnostic measures. Dr. Jane R. Mercer, a sociologist at the University of California at Riverside, has studied the effects of such labeling. She has been concerned with the relationship between sociocultural factors and mental retardation, and especially the fact that children from minority groups are more often labeled as retarded than are children from the white middle-class culture. Once labeled as retarded on the basis of a test score, the children are treated differently. They are put into special classes from which it is almost impossible for them to move into a regular class. Other children, and even their own parents, may act differently toward them. Their lives are often totally changed as a result of the label placed on them—which may or may not be accurate.

This NIMH program report describes the research that Dr. Mercer has been doing. It includes a description of the "system of multicultural pluralistic assessment"—an instrument for evaluating the abilities and potentialities of children based on many different aspects of their performance. The report is intended for a wide audience: Mental health professionals conducting research on mental health of children, child psychiatrists and psychologists, teachers, parents—all will find information of relevance to them in this report.

Bertram S. Brown, M.D.
Director
National Institute of Mental Health

Labeling the Children

We are often told that America's greatest natural resources are the potentialities of its young people. But what if we discovered that the development of intellectual abilities for large numbers of our young people was being systematically stultified and inhibited? And that the main instruments being used to this effect were the very tools that were supposed to pick out and reveal the able and the gifted—the intelligence tests given to all schoolchildren?

Take the case of Anthony, a 5-year-old Mexican American child living in Southern California with his parents and six brothers and sisters, who generally speak only Spanish at home. His mother speaks nothing else. Despite this handicap, he made a score of 105 on the Wechsler Intelligence Score for Children—Revised (WISC-R), which is in English and is generally considered the best all-around intelligence test used for young children.

This score puts him a little above the norm, an acceptable result. But if children from his own particular sociocultural background are singled out, their average score is found to be 83.5; and Anthony is 21 points above that, a high score indeed. Therefore, if adjustments were made for cultural and language differences and for the unfamiliarity of children with Anthony's background with the American core culture, he should rank in the top 5 percent in intelligence for all children—a truly gifted child, capable of doing well in college and of contributing a good deal to American society. But if the school should keep him in the track set for him by the WISC-R score alone, a little above average, then he will move along the relatively slow path of the tens of millions of

"ordinary" children, perhaps never realize his abilities, probably never go to college.

Or take the case of Allen, a 15-year-old black child who has been in special education classes since he was 8, because of a score of 78. Unmotivated in that class, he has never learned to read. Merely being in it almost all of his school years has effectively branded him a retardate in the eyes of his peers and apparently also of his teachers. His mother speaks of his situation with resentment. "Let's face it, children can be real cruel. It's as if [the youngsters in those classes] have a sign around their neck for everyone to read. Many of the teachers talk down . . . ." As a result, Allen is "a rebellious teenager." From the path he is on and has been treading for 7 years, it does not seem hard to predict what his future will be: a black youth, unable to read, a dropout or a graduate with a worthless diploma, frustrated and angry, branded "mentally retarded." Yet his mother says of his abilities away from school, "Around the neighborhood he's really a handyman. He fixes electrical appliances, moves things, mows lawns. He'll repair a lamp or something that's broken, if he wants to."

Since she started to write her doctoral dissertation in the early 1960's, Dr. Jane R. Mercer has been watching this labeling process and its consequences—at first with concern, and then with increasing alarm. As a sociologist she had never thought to make retardation her main research interest—like most of the rest of us she assumed that it was primarily a medical problem. But the subject for her doctoral dissertation was "An Analysis of Factors in the Family's Withdrawal of a Patient from a Hospital for the Mentally Retarded." She found that the disadvantaged, the blacks, the Chicanos fought much harder than the "Anglos" to get their relatives out of those hospitals; and when they did get them released, the so-called retardates usually disappeared quickly into their own social and racial milieus, worked and functioned well, and generally were little different from those around them. "Their relatives never for a minute thought they were really retarded. They simply assumed that this was another thing white society was doing to them, and they had to fight it through."

In 1963, partly because of this experience, Dr. Mercer was appointed field director of an 8-year research project studying the epidemiology of mental retardation in Riverside, California, a community not far from Los Angeles. This study and the conclusions, insights, and compensatory programs that have resulted from it have now become world famous and are having an increasing influence on considerations of the validity of intelligence tests particularly as they affect minority groups (and the conclusions about racial in-

feriority that had been deduced from their scores), and on curricula and school philosophy in areas where there are substantial numbers of "non-Anglo" core students. Her criticism of IQ scores and tests for being culture-bound *achievement* measures and for being grossly discriminatory and damaging—ideas that would have brought a strong counterattack two decades ago—are now being widely considered and often accepted. "Retardation is a sociologist's concern," says Dr. Mercer. "We are dealing with assigned social roles and the process of labeling. Only with a very small percentage do we deal with recognized medical problems. Labeling is a social process which can be very damaging if it leads to dead-end educational programs and reduced life options."

What happens to the disadvantaged children who have been falsely labeled? "We call it 6-hour—or situational—retardation," says Dr. Mercer. "Only during school hours are most of these children considered retarded." Outside of school or during vacations most of them take their full places in their communities. But, at least a high school diploma and some specialized training are needed to gain access to society's resources, and it is harder for such children to finish school and get training. "To label a child, to assign a child to a derogatory social role is dangerous because it may result in a self-fulfilling prophecy." First the chief institution of labeling, the school, has its own means to enforce its judgment—the tracking system. Once children are put in the lowest track, there is little chance they will rise out of it; once they are put in special classes, there is even less chance they will return to the regular program. One black mother in Dr. Mercer's study says of her child in a special class for the educable mentally retarded: "The only reason he consents to go is because we have promised that he'll be taken out of that EMR class. The teachers have asked us to let them put another one of our kids in EMR. We said an emphatic 'no!' because we knew what it was all about." Second, of course, the child or youth may come to believe the label himself—and if he doesn't, there are plenty in a competitive society to remind him.

Non-English-speaking or nonwhite parents generally tend, after some bewilderment and exposure, to learn what it is "all about." The Caucasian parents who are closer to the core culture come to accept the judgments of the schools. But then their children do better on the tests. More are labeled superior; fewer, retarded.

The 8-Year Riverside Study

The 8-year study of Riverside, California, was initially conducted at Pacific State Hospital and codirected by the hos-

pital's director, Dr. George Tarjan, and the late Dr. Harvey F. Dingman. Says Dr. Mercer, "We were to study the epidemiology of the community of Riverside to determine the prevalence rates of mental retardation. Seldom, if ever, had such a large-scale, long-term study, covering so much of an entire city of 80,000 been undertaken. We used two approaches. On the one hand, we did a household survey of a representative 9-percent sample of the total population. We screened about 7,000 persons under 50 for signs of retardation. To keep our data consistent and reliable, we did not try to find out and use the criteria actually used by the schools or other agencies. Instead we used the official bipolar definition of the American Association for Mental Deficiency (AAMD). Both subnormal IQ and subnormal adaptive behavior had to be present."

The second approach was to contact all the community agencies that served the retarded and develop a case register of those persons who had been so designated or labeled. Dr. Mercer and her team contacted over 241 agencies and eventually developed a list of 812 persons.

Because they were "studying the labeling process," they did not impose a standard definition on those considered retarded by the organizations, but asked each agency to use their usual definitions. There were many duplications; the 812 names had been nominated 1,493 times. However, when the organizations were grouped into different categories, the investigators found that the public schools had contributed 429 persons—more than half—to the register, 340 of whom had not been nominated by any other organization. The next most frequent contributors were the medical facilities, with 166. "When we studied those jointly nominated by various organizations," says Dr. Mercer, "the public schools clearly held the commanding position. They not only labeled more persons, but they shared their labels more widely throughout the whole town."

What kinds of information did these different agencies use to determine labels that might remain on the children for life? "We found that the public schools rely primarily on IQ test scores." Almost all—99 percent—of their cases had been given intelligence tests; but only 13 percent had received, or been referred for, medical diagnosis that might have identified the reasons for low IQ's and suggested appropriate treatment. Law enforcement agencies also relied primarily on IQ tests. Neither measured adaptive behavior except perhaps in the most subjective way.

The California Department of Mental Hygiene, on the other hand, used both IQ scores and medical diagnoses for over 90 percent of its nominees.

4

What level of IQ determined who was labeled and who was not—that vague but extremely important borderline? The standards were not consistent. Reports Dr. Mercer: "When we first started the study, we discovered that there were three different cutoff levels used by people in the field. The AAMD recommended that anyone more than one standard deviation below the mean be considered subnormal. That includes about 16 percent, all those scoring below 85 on an IQ test." The public schools, on the other hand, were mostly using a cutoff of IQ 79. "In the textbooks written by the people who developed the tests, men such as Wechsler and Terman, a person was defined as subnormal only if he was in the bottom 3 percent of the population, or two standard deviations below the mean, an IQ of 69 and below." The Report of the President's Panel on Mental Retardation also puts the line at the bottom 3 percent. Nevertheless, "we found all three of these definitions being used simultaneously." Dr. Mercer found nothing in the literature that discussed which of these divergent cutoff lines was the proper one. No one even seemed to address the issue, thus compounding the probability of erroneous labeling. Even using the same cutoff, a difference of one or two IQ points (which might have resulted from an error in scoring, or be within the normal range of random statistical error) might determine whether a child was labeled mentally retarded or normal. With a variety of cutoff levels all operating simultaneously, problems of erroneous labeling are multiplied.

But this was the situation. Dr. Mercer and her colleagues concluded that the 3-percent cutoff was preferable because it was the one "least likely to label a person mentally retarded who was regarded as normal by persons in his family and neighborhood." (If subnormal behavior is added to the criterion of subnormal IQ, that 3 percent shrinks to about 1 percent.)

But the public school net, which caught the most children, was much wider: 46 percent of those nominated by the schools for the retarded list had IQ's above 70, and 62 percent had no reported physical disabilities that might have explained some of that retardation.

To fully understand the importance of the role of the public schools in the labeling process, not only for themselves but for all the community, let us consider some facts revealed by the study:

1. The public schools were the primary labelers. They labeled more children than any agency; and they spoke with great authority.
2. They were the pivotal screening and referral agency for the entire community. They shared their labels and testing ma-

terials widely. Other organizations received psychological reports from the schools and used them. They also referred children to the schools for assessment.

3. The public schools relied mainly—often exclusively—on IQ test scores to make diagnoses of mental retardation. As noted earlier, 99 percent of those nominated by the schools had been given an IQ test while only 13 percent had received medical examinations. Of those nominated by other agencies (except law enforcement), 50 percent had received medical checks.

4. The norms of the school were therefore clearly embodied in whatever is measured by standardized tests, particularly the Stanford-Binet or Wechsler Intelligence Scales. Any kind of bias in the tests themselves could obviously have serious consequences.

5. Children labeled as mentally retarded by the schools were less deviant than those labeled by other agencies.

6. Labeling was done mostly during the elementary grades. Eighty-three percent of the "retarded" were classified, referred, and placed during their elementary school years.

7. Who, specifically, classified and referred them? Chiefly the teachers. For instance, during one school year, 71 children were placed in classes for the mentally retarded. All but two were first identified and referred by a classroom teacher. Not one child was referred by a parent or relative. Dr. Mercer and her associates examined the records of all the children in EMR classes throughout the district at the time of the study. They found that 90 percent were referred by classroom teachers, 3 percent were transferred in from other districts, 5 percent were referred by the probation office or other sources, and only 2 percent were referred by a parent or relative.

8. Those referred as possibly mentally retarded by teachers were tested more readily by psychologists. In other words, psychologists relied heavily on teacher judgments and decisions. Ninety percent of those referred for "mental retardation" were given IQ tests by the psychologists compared to 70 percent of those referred for "academic problems"; among those referred for "behavior" problems, only 40 percent were examined.

9. The Catholic parochial schools labeled only two among their 2,800 children as mentally retarded. Although the investigators hypothesized that parochial schools might tend to discharge the most troublesome students to the public school system, no evidence was found to support this hypothesis. The primary reasons were that the Catholic schools did not use standardized tests or employ psychologists to do the testing, nor did they (at the time of study) have special classes to serve the mentally

retarded. They were not systematically looking for mental retardates. On the other hand, Dr. Mercer's group did find 1.1 percent of the Catholic school population who, when tested by project psychologists, scored below 79 on the Stanford-Binet and would in all probability have been referred for classes for the mentally retarded had they been in public school. But the parochial schools did not single them out.

The overall process is evident. Says Dr. Mercer, "The amount of labeling done by a social institution is related to institutional structures and testing procedures." There are no automatic, "objective" measures; in the last analysis somebody in some institution makes a more or less arbitrary decision that a child's behavior does not meet his or her expectations.

10. Black and Spanish-surnamed children tested by the school psychologists were far more likely to score 79 or below on the intelligence tests than Anglo children. There were four times as many Chicanos labeled retarded, and twice as many blacks. They failed the tests in higher numbers; therefore they wound up on the lists of labeled retardates.

11. Even among those with scores below 79, the poor, those with Spanish surnames, and those the teachers thought retarded were most likely to be placed in special classes. Those recommended for placement made up only 64 percent of those with scores below 79. Their IQ scores were not significantly lower than the rest.

12. What happened to the "MR" children themselves? Only 19 percent returned to regular mainstream classrooms. Almost a quarter—23 percent—dropped out of school without adequate education; 46 percent were eventually kicked out of school or placed in other programs or institutions; and 12 percent aged out of the programs. For four-fifths of the children, entrance into the special classes was permanent removal from the educational mainstream.

13. School-age children were labeled MR much more often than preschoolers or adults, who were consistently underlabeled. This once more reflects the role of the school and of teachers as chief labelers. Among preschoolers only, those with the most physical disabilities and the lowest IQ scores were labeled. After graduation or departure from school, only the most physically and mentally subnormal adults were classified by agencies or neighbors as retardates. The rest of the adults simply disappeared into the general population and got by.

14. The poor were consistently overlabeled; the well-to-do underlabeled, compared to their percentages of the population. These disproportions persisted within the ethnic groups themselves—well-to-do Chicanos had a smaller proportion of MR's than poor Chicanos, etc.

15. Children labeled MR from the more well-to-do were more likely to have physical disabilities than those labeled mentally retarded from lower socioeconomic levels. High status children were more likely to have been defined as retarded by parents during the preschool years. The poorer children defined as MR were more likely to be physically indistinguishable from other children, were not as likely to have been considered retarded by parents, and were more likely to have been first so classified by a teacher. In other words, although fewer well-to-do children were classified as subnormal, those who were had obvious and physically distinct characteristics.

16. Children from poorer homes and neighborhoods were more likely to be considered retarded only during school hours, or for purposes of schooling, than more prosperous children. The poorer children were more often "6-hour retardates." Seventy percent of those considered retarded in school were not perceived as such by family or neighbors—or, for that matter, other agencies.

17. On the other hand, lay persons, not having the advantage of the standardized tests, tended to label only the extreme cases —the most physically and mentally subnormal—retarded. Those labeled by lay persons in the field survey had average IQ's of 55, while those labeled by the schools had average scores of 67.4. Over half—52 percent—of those nominated by lay respondents had physical disabilities, compared to 38 percent of the school candidates.

18. Among those without physical disabilities, proportionately more males than females were labeled MR. Among those with physical disabilities, the percentages were equal. It is not clear why this is true. Girls tend to be more docile, to do less acting out. In the early grades girls often seem to fit into school routines better. And, of course, the great majority of teachers who come in contact with these children and do the evaluating are women. But these are speculations.

What Price Behavior?

By the end of the study, Dr. Mercer and her associates found themselves with three overall conclusions that could not be escaped and that had strong implications for future programs and policy.

First, though the AAMD had called for a bipolar definition of mental retardation, in which the factor of adaptive behavior was just as important as IQ score, the clinicians, psychologists, and physicians who were examining the children sent to them by the teachers were simply not measuring or including adaptive behavior. Most diagnoses of MR were being made by these experts almost entirely on the basis of

the IQ score alone. Some physicians were giving medical examinations, and in some cases they found physical anomalies. However, in most cases even the physical checkup was absent.

To leave out the adaptive behavior test is a very significant omission. As we have seen, the majority of those labeled retarded were not considered subnormal by those who knew them best away from school. Laymen, who do not give intelligence tests, considered very few of them subnormal. This finding is the more remarkable because, in most cases of mental or emotional deviance, behavior is the primary symptom that alerts the officials—and behavior is often the only test. "She acts funny"; "he's overly aggressive and noisy"; "she withdraws, and sits by herself"—are typical of the factors that usually attract attention and referral. Only later, if then, does testing occur. Mental patients are judged to be improved or well, in most cases, by their behavior. The great exception therefore seems to be the school-labeled mental retardates without physical anomalies.

Why weren't the therapists measuring adaptive behavior? Partly perhaps because the IQ tests over the years have built up a mystique: They are reliable and objective; they measure potential to live and master complicated material in a necessarily complex society. Everything else is treated as "subjective."

But Dr. Mercer found another reason, more directly to the point. They could not measure adaptive behavior even if they wanted to because they were no scales available for them to use. Intelligence tests have been refined over many years. Adaptive behavior has no comparable measures. She says, "We concluded that such a scale needed to be developed as soon as possible, and that clinicians should begin to systematically measure adaptive behavior as well as IQ scores."

Their second conclusion was that the various cutoff levels now being used simultaneously to determine the border between retardation and normality would have to be discarded and replaced by a single standard if the designation was to have any meaning. As we have seen, Dr. Mercer and her associates adopted the 3-percent cutoff as the least likely to stigmatize persons who would be regarded as normal by family and friends. She reports: "In their new nomenclature, the AAMD now uses that 3-percent cutoff. They had been using 16 percent at the time of our study; but not any longer."

We have already mentioned the third conclusion: "We rediscovered and were greatly impressed by something already generally known. Tremendous cultural biases exist in the IQ test."

Old or not, this conclusion is likely to cause the greatest controversy and have the greatest long-term impact. First,

it is in an area already saturated with controversy, strong feelings, and allegations. Second, probably never before have biases in IQ tests been documented with such a wealth of data, nor have the consequences of those biases been more specifically outlined. "The test," says Dr. Mercer, "is not appropriate when used with many of the poor, and is certainly not appropriate when used with most Chicanos and blacks. It simply is not a true measure of their biological potential because they did not have the same opportunities to learn the cultural material that is used in the test."

IQ has come to mean, to most people, the measure of innate or genetic intellectual potential. To them it is something apart from schooling or the environment in which the child grows up, although a child's potential might not come to be fully realized without proper schooling or encouragement.

Over the years it has been taken for granted that intelligence tests were more or less accurate measures of intellectual potential—particularly the newer tests which have been carefully refined and revised according to accumulated experience and knowledge. Measuring "potential" was not their initial purpose. Binet's original test in 1905 was designed simply to identify those children who would not succeed academically so that they could be placed in schools for the subnormal. The test was to be a measure of projected performance in school.

Dr. Mercer has emphatic opinions on this point. In cases of racially homogeneous suburbs where those taking the tests have similar backgrounds, homes, and cultural experiences, intelligence tests such as the WISC-R may well show relative innate intellectual capacity. "But in all other cases," she says, "particularly in racially mixed student groups, it's just another achievement test, one of many the kids are forced to take. It measures the achievement of mastering the Anglo-core culture that dominates the schools."

The Anglo-core culture depends on English not only as the primary language, but for most practical purposes as the only language. The ability of children to speak English is essential and is a basic test of ability to perform in the school. If they cannot speak English well, they have a deficiency that must be corrected as soon as possible. In the meantime in their studies—and in their intelligence tests—they will be marked down.

The values and procedures of Anglicization were established by the early settlers as the standard. Later immigrants were expected to conform to it. Usually, after at least one lost generation, which often saw its speech and customs dismissed or denigrated, the majority did conform, particularly

the whites from European countries. Anglicization became thoroughly established as the norm by which all might be measured, particularly in the schools. Anglicization not only includes language but cultural values—"those," says Dr. Mercer, "of Benjamin Franklin and the Puritan ethic." She likes to quote from Clarence J. Karier that they include "a fear of poverty, a desire for wealth, a respect for work, a need for thrift, and a keen sense of duty to be a productive useful citizen of the community." To these we might add traditional American middle-class values closely associated with mainstream teaching: motivation to excel and to win, the ability to delay gratification (and curb present appetites) for future rewards, belief in the future if one works hard now, and belief in material rewards as the measure of success. The intelligence tests, like the schools, tend to reflect all these. And they also emphasize another: speed as a function of excellence—both achieved and innate. They assume that the bright child not only gets more answers right, but finishes first. If he is slow, even if it is because he would like to explore all possibilities, he may lose the race.

Willingness to best others in competition is essential. The winning scores are used as the bases of many rewards: higher tracks in school, enriched curricula, greater expectations from teachers (who usually adjust their time and efforts accordingly), and perhaps eventually scholarships.

These values, and the English language itself, may be only imperfectly understood or accepted by those who speak foreign dialects or live in the ghettos. In fact, many middle-class white children accept them only as part of the game they are expected to play to get through school. But by this yardstick all will be measured.

Since Anglo conformity assumes one central culture, it also assumes only one normal curve by which the behavior of all children can be classified, regardless of background. Such a statistical model generates its own abnormals. If there is any variability in the population at all—and in the American population there is plenty—2.5 percent will always be abnormally high and the same percentage abnormally low. The means may change—intelligence test scores have gone up more than 1¼ standard deviations since World War I—and thus yesterday's "low normals" become today's "mental retardates"—but the stigma hardly changes.

Sociocultural characteristics not only accounted for the differences in IQ test scores between ethnic groups, but for many of the differences within each group. Stepwise multiple-regression analyses were run for Chicanos and blacks in Riverside according to the degree each subgroup approached white lifestyles and standards. The Chicanos and black children with

the highest IQ's tended to come from homes with fewer children, both parents in residence and with some education and determination that their children would get more, and a breadwinner with a relatively high-level job (usually white collar). More of the families lived in their own homes. In short, the closer to the white community norm, the higher the scores on the WISC.

"Since the tests are so culture-specific," says Dr. Mercer, "and their extensive use had caused so much damage to minorities and the poor, our third conclusion was that a system needed to be developed that would compensate for these biases. We proposed that the intelligence test be made a more valid measure of those who are not middle class and Anglo, to give them a more fair assessment." In 1970, Dr. Mercer and her associates received a grant from the National Institute of Mental Health to develop such an instrument and to take the initial steps to put it to practical use. It was to be called a System of Multicultural Pluralistic Assessment—SOMPA.

Says Dr. Mercer, "We call it pluralistic assessment because in evaluating a child many different aspects of his performance should be taken into account. No single measure—such as the IQ—should be the determining factor." The pluralistic measures of SOMPA were not designed as laboratory tools but as practical screening instruments to be used by schools and clinics. They have been standardized on a representative sample of the public school children of California—2,100 pupils and their mothers, one-third each black, Spanish-surnamed, and Anglo. "Our battery is designed for the early elementary years, for children 5 to 11. Those are the critical years, the ones in which children are usually diagnosed, assessed, and labeled."

The Three Faces of SOMPA

"SOMPA uses three conceptual models for assessment," says Dr. Mercer. "That requires the user to have greater cognitive sophistication than the usual single model approach. We feel that at least three perspectives are needed to understand children who come from so many different places in America's sociocultural smorgasbord." These assessment paradigms—the medical model, the social system model, and the pluralistic model—have differing premises and approaches. Each presents a different perspective on the child's behavior. But Dr. Mercer views plotting the position of the child by use of the three measures as a form of triangulation, starting from different directions, rather than a search for agreement.

Each model has a distinct purpose, a distinct definition of normal behavior. Each creates a different picture of the child. Each provides useful and unique information.

The *Medical Model* is perhaps the most powerful and pervasive model used to describe the range of human variety, and certainly of human abnormality, in existence.

In this model normality is basically a negative quality—the absence of disorder or of pathological symptoms. Normality is health—abnormality is illness; health is good, illness is bad. When we say we feel physically "well," we do not really mean that we feel anything much (except perhaps a vague sense of well-being). We feel things when we are ill—pain, nausea, dizziness, and so on. Although, of course, we may feel well and still be ill—a condition that might make us suspicious and cautious. But even then an expert will be able to find biological abnormalities.

In the medical model malfunctions are abnormal, whether they are physical or behavioral. When the medical model is used, the inference is usually made that undesirable behavior has some physical base.

Another assumption of the medical model is that the flaw, however it originated, now lies in the patient and can be treated by treating the patient directly and individually. The sociocultural characteristics are not directly producing the symptom; so, for purposes of treatment, and in large part of diagnosis, they are irrelevant. Appendicitis is treated the same no matter what language the person speaks or where he lives. Background is relevant only when it affects the disorder—as when cold slum houses aggravate tuberculosis. Therefore, the medical model is not culture bound.

The medical model focuses on deficits—basically it is concerned only with the sick. It seldom defines degrees of health—the healthy are usually lumped together. It searches for fairly simple and direct cause-and-effect relationships: A germ causes a disorder. If the cause is known, the treatment cannot be far behind and can be directly applied. Social and cultural factors tend to be much more subtle and complex in their interactions and are usually discounted.

There is a final characteristic of the medical model that is extremely important and, when applied to a situation in which that model is not appropriate, can cause considerable harm. A basic premise is that health is not merely different from illness, but a far superior condition. Moreover, as noted, illness can often be hidden—cancer and hypertension are often secret, painless, and invisible killers. Even in behavioral changes, a brain lesion or chemical imbalance may be responsible. Therefore, it pays to assume the possibility of trouble, if there is any indication at all, and to search for it. It is far better to be safe than sorry. Labeling is not a real issue. If a physician feels a child might have a small tumor, it is far better to assume that he does until it can be verified or dismissed. This classification will do him no real harm; it may save his life.

Preventive medicine is based on this premise. But labeling a child for a behavioral or social flaw or deviancy may be a very different thing; and the medical model should not be used outside its proper context.

"Since it's so serious to overlook a biological malfunction in a child, we recommend in the SOMPA that those one standard deviation below the mean—within the lowest 16 percent on the medical measures used—be referred for an indepth medical examination by a medical practitioner," says Dr. Mercer.

"One of the things we have found is that most schools simply do not have adequate medical records for their children and that most children have not had adequate checkups—at least it's not in the school records. One good effect of the SOMPA will be to get children medical care who never had it before."

Six of the measures used by SOMPA are interpreted, and fit, within the medical model: the Physical Dexterity tasks, the Bender-Gestalt test for young children, the Health History inventories, and tests of vision, hearing, and weight-by-height. Note that the WISC-R is not listed as a medical model measure because it does not meet the assumptions of that model and cannot be used effectively across cultural boundaries.

The *Social System Model* says that abnormality is deviance—behavior that violates accepted standards. But deviance depends on role—what is deviant in one situation or role may not be at all deviant in another. For instance, it is deviant behavior to try to charge into someone and knock him down in the classroom, but not on the football field. There are as many definitions of deviance as there are roles. Definitions of abnormality are not only bound by the values of the social system in which they take place, but to the role in that system. There may be as many roles in a child's everyday life as in a Shakespearean play: son in the family, fifth-grade student in school, first baseman on the team, leader of the gang, treasurer of the club, and so on. Each is a role; each is in a separate social system.

Unlike the medical model, there are degrees of normal in the social system model (some people are better first basemen than others), the distribution of scores is normal instead of being flattened at the top, and the ceilings are high. Since the norms of each system define what is normal or abnormal, the assessor must know something about the norms of that group. A single blanket standard applied by someone from the outside who may be using different norms is meaningless. An outsider can gain some understanding of the norms by asking group members to describe the way other members should perform. Says Dr. Mercer, "This is essentially the approach we used when we talked to the mothers from different backgrounds. What were the behavior and the roles of their children at home,

in the neighborhood, the school, and the community? We used this information to develop the Adaptive Behavior Inventory for Children (ABIC), one of the SOMPA social system measures."

Assuming abnormality under the social system model until proof is forthcoming can be dangerous since it tends to label the child and to create the possibility of developing a self-fulfilling prophecy.

The assumptions of the social system model are quite different from those of the medical model. Normality is not a trait of the child, but simply a definition of performance in some social system. The norms are not absolute. They are determined in a political process and may change with people and time. Those with the most power can impose their definitions of "normal" on others and make those definitions binding in anyone who wants to continue to participate in the system. If you want to play in the ball game, you have to obey the rules laid down by those who have the ball and bat; if you want to keep your job, you seldom oppose the boss. The example with the most immediate relevance, of course, is the school system. There the power lies in the bureaucracy and the professionals who run it; and, as noted, they administer and teach according to the norms of the white, middle-class Anglo culture.

Cause-and-effect reasoning is not characteristic of the social system model. Deviance is not biologically determined. And there is no such thing as unrecognized social deviance—there are no covert misbehaviors; deviance is by definition a form of overt behavior that has been evaluated by others according to some social norm. If a child has heretical thoughts about his social system, they are not deviant if he does not act on them.

In SOMPA there are seven measures within the social system model: the WISC-R using standard norms and the six scales of the Adaptive Behavior Inventory for Children. The WISC-R is used to measure the child's School Functioning Level (SFL). The ABIC, using the mother as informant, measures the child's role performance in the family, the peer group, the community, nonacademic school roles, earner and consumer roles, and self-maintenance.

The *Pluralistic Model* assumes that all tests measure learning. Subnormality is low performance on a test in comparison with the scores of others from similar backgrounds. It assumes therefore that all children from the same sociocultural environments have about the same opportunities to learn material and have had similar experiences with taking tests. In short, when social, cultural, and emotional factors are equal, those children who learn and perform best have the greatest learning potential. Differences in scores earned by ethnic and

culturally distinct groups are due to biases in the tests themselves. The pluralistic model takes it for granted that we are a Nation of peoples, rather than a single homogeneous group that can be plotted on a smooth, normal curve.

The problem then becomes how to determine a child's Estimated Learning Potential (ELP) in spite of cultural differences, when the tests reflect and are specific to a single culture. The usual method has been to attempt to construct a culture-free test that will measure the potential, regardless of background. So far all such attempts have been unsuccessful. Perhaps a culture-free test is a contradiction in terms. All tests must use some language or points of reference, and these associations will not be equally familiar to all children.

The pluralistic model does not try to change the content, language, or procedures of the test used—the WISC-R. What it does instead is to use multiple-regression equations to determine the average score on that test of children from closely similar backgrounds. This average then becomes the norm to which the child is compared and from which his ELP is determined, rather than the "100 IQ normal" of the test itself, which is based on the performance of children from the dominant culture. The case of Anthony, the 5-year-old Mexican boy, is a good example. He scored 105 on the WISC-R. But the norm for children from his own background is 83.5; so he is actually 21 points above the norm rather than 5 points above the normal. In short, he has a very high ELP that was not directly evident from the WISC-R. It is on that Estimated Learning Potential that he should be assessed and placed, not the standard IQ.

Incidentally, it might be noted that 83.5 is not very far from the "retarded" score of 79 used by many schools. But the Adaptive Behavior Inventory for these children does not show that the large percentage who fall below 79 are retarded.

Obviously the pluralistic method of rating normal and abnormal is culture bound. Obviously also it is an asset model: It tends to uncover potential in children that might have been masked by cultural difference, as with Anthony. Once the differing norms are taken into account, the pluralistic model has a full range of scores to which any child can be compared. There are many degrees of normal and abnormal.

There is one assumption about the pluralistic model that seems obvious and simple, but has great and controversial significance. "We have assumed," says Dr. Mercer, "that the differences in learning potential between individuals are considerable, but differences between sociocultural groups are negligible." Put bluntly, the pluralistic model assumes that all races and nationalities are created equal in potential. There

are no genetically superior races—at least not in the intellectual sense.

The great majority of scientists accept this view as the most decent, the most logical, and the most in accord with available facts. To assume anything else without incontrovertible evidence, they believe, is to enter a scientific wilderness and, worse, to do great harm by reinforcing prejudice and discrimination. From the great mass of epidemiological data she has accumulated, and in order to be able to set up a reliable mechanism that could compensate for obvious inequities and mislabeling, Dr. Mercer frankly felt she could reach no other conclusion.

Nevertheless, until intelligence and other tests of potential are truly culture free and accurate, the equality of learning potential across race and culture must be accepted as an article of faith. Blacks, Chicanos, the poor, and other minority and disadvantaged groups do by and large make lower scores on tests. They do not usually function as well as the dominant Anglos in the mainstream culture. They do not do as well, on the average, economically as the white middle-class—to many the ultimate test of competence. Dr. Mercer and her associates, by taking this stand as emphatically as they have and by creating SOMPA as a means of correcting for cross-cultural exposure to test materials and of equalizing opportunity, have come into strong conflict with those who consider the tests to be at least moderately accurate. They have become almost as famous for their defense of equality as for the invention of SOMPA.

In SOMPA, the WISC-R has been used as the instrument to develop pluralistic norms. In view of Dr. Mercer's frequent strictures against intelligence tests, this reliance might seem surprising. But the reasons are simple. The WISC-R is the most widely used, individually administered test in the United States; consequently it is the most important assessment tool used for placement and planning. It has recently been revised and renormed. It is the best predictor of school performance. The final reason is the most important and the most humane. Says Dr. Mercer, "We believe that the whole point of educational assessment is to identify children who may fail in school as early as possible in their school careers—not after they have failed again and again, and become labeled, discouraged, and alienated.

"The achievement test used therefore must be individually administered, cannot depend on reading and writing skills, and must be a good and early predictor of what school performance will be like." The WISC-R meets these criteria.

Dr. Mercer does not consider the three models, or the measures used by each, to be interchangeable or even additive.

She feels that all measures should be given and the results plotted. "All measures have been standardized for profiling on three sets of profiles: for the medical model, the social system model, and the pluralistic model." The measures based on the models are then used much as a navigator or mathematician might use triangulation of three apparently unrelated directions to find the precise spot of intersection, where the target, or the child, might be. "When using all three models," she says, "if findings are similar, we can make interpretations with a high level of confidence. When findings differ, we have to be much more cautious about reaching premature judgments and use individualized programing." So the SOMPA profiles can provide a basis for deciding what treatments or compensatory programs to use in different areas, as well as determining what the child's real potentials and strong points are.

SOMPA Faces the Future

SOMPA was designed to be used, and it is being used, in some schools and clinics, as well as in Dr. Mercer's laboratories. What of the future? Says Dr. Mercer, "We are working up some profiles of different kinds of children on the basis of various SOMPA batteries." One set of children that could be readily identified on these profiles are those with major organic damage. "These children might or might not have adequate adaptive behavior. But they would be low on the physical dexterity battery, have suspicious health histories, and have difficulty with the Bender. They therefore should be spotted and referred early for medical diagnosis to determine if there is any biological damage."

Another obvious type of child whose SOMPA profile will become familiar is the child who is culturally different, but has no other apparent problems. Dr. Mercer says, "These children would score low on the standard norms for the WISC-R, but when the culturally appropriate norms are used they would be in the normal range. They would have normal health histories, normal dexterity, normal behavior among their peers and at home. Their academic problems result from cultural difference, and it is in this area that help would be given in the school."

Other profiles and general classifications of children will emerge. Once these have been identified statistically, it will be possible to sketch out, in rough form at least, the interventions each group would need. "For children who do not do well in the intelligence tests because their backgrounds have not prepared them for the associations and the value systems of the core culture," says Dr. Mercer, "we will be able to

provide programs much more appropriate than placing them into classes for the retarded."

SOMPA and the research that led up to it have been well received, overall, because in a very real sense they are children of their time. The increasing criticism of intelligence tests, the civil rights struggle, and other events have changed the atmosphere of only two decades ago.

Says Dr. Mercer, "Most of the conclusions reached at Riverside were neither new nor startling. They too argued that existing tests underestimated the abilities of working-class children, and proposed a series of gamelike tasks they believed were culturally fair. But, according to reports of the time, their work was vehemently attacked by outraged psychologists and was soon abandoned because it did not predict success in school." The testers were in public favor then, and the cause of better schooling for the poor did not have today's militant defenders.

But when the findings from the Riverside epidemiology study were first reported publicly in 1965, segregation in schools had been illegal for 11 years, and Title VI of the Civil Rights Act of 1964 had prohibited discrimination of any person because of race, color, or national origin in any program receiving Federal money.

Unlike many of the children it was designed to help, SOMPA was conceived, was born, and has grown through its tenderest years in an environment of increasing acceptance and understanding. In California as well as in much of the rest of the Nation, successive court decisions and legislative acts have gradually brought to public attention, and provided support for, six issues that are basic to the proper assessment, classification, and placement of children in the public schools:

1. The requirement that testing must take into account a child's language and cultural and racial origins and be nondiscriminatory;

2. The requirement that a child's adaptive behavior must be evaluated before he or she is placed in a special class for remedial or supplementary instruction;

3. The child's right to a place in the mainstream of education, instead of being shunted off into lower tracks or special classes;

4. The child's right to be tested in his or her primary language, whether or not it is English;

5. Better definition of the right cutoff level for subnormality, or for placing children elsewhere than in the mainstream;

6. The right of the child and of parents to due process and participation when decisions about placement or other spe-

cial treatment are being made—in other words, no arbitrary or "professionals-know-best" treatment without appeal or consultation.

"We must come to accept the fact that America is really a pluralistic society." says Dr. Mercer. What may be at least as important is that we come to understand that segregation and punishment of children because of social and cultural differences, whatever the process or rationale, are no longer acceptable or tolerable in the United States. And assessment systems like SOMPA must be made and used now, so that no Anthonys are denied college and no future Allens will be left embittered and unable to read.

Supplementary Readings

Mercer, Jane R. Social system perspective and clinical perspective: Frames of reference for understanding career patterns of persons labeled as mentally retarded. *Social Problems*, 13(1): 18–34, 1965. Reprinted in: Weinberg, Martin S., ed. *Deviance: The Interactionist Perspective*. New York: Macmillan, 1967.

Mercer, Jane R. Institutionalized Anglocentrism: Labeling mental retardates in the public schools. In: Orleans, Peter, and Russell, William, Jr., eds. *Race, Change, and Urban Society*. Urban Affairs Annual Review. Vol. 5. Los Angeles: Sage Publications, Inc., 1971.

Mercer, Jane R. Pluralistic diagnosis in the evaluation of black and Chicano children: A procedure for taking sociocultural variables into account in clinical assessment. *California School Psychology*. 19(1): 23–28, 1971. Reprinted in: Wagner N. N., et al., eds. *Chicanos: Social and Psychological Perspectives*. 2nd ed. St. Louis: C. V. Mosby Co., 1975.

Mercer, Jane R. Who is normal? Two perspectives on mild mental retardation. In: Jaco, E. G., ed. *Patients, Physicians and Illness*. Glencoe, Ill.: Free Press, 1972.

Mercer, Jane R. *Labeling the Mentally Retarded*. Berkeley: University of California Press, 1973.

Mercer, Jane R. The myth of the 3% prevalence. In: Tarjan, G. et al., eds. *Sociobehavioral Studies on Mental Retardation: In Honor of Harvey F. Dingman*. American Association on Mental Deficiency Monograph Series, 1, 1973. pp. 1–18.

Mercer, Jane R. Latent functions of intelligence testing in the public schools. In: Miller, L. P., ed. *The Testing of Black Students: A Symposium for the American Educational Research Association*. Englewood Cliffs, N.J.: Prentice-Hall, 1974. pp. 77–94.

Mercer, Jane R. A policy statement on assessment procedures and the rights of children. *Harvard Educational Review*. 44(1):125–141, 1974. Reprinted in: Chess, Stella, and Thomas, Alexander, eds. *Annual Progress in Child Psychiatry and Child Development*. Vol 8. New York University, 1974.

Mercer, Jane R. Crosscultural evaluation of exceptionality. In: Meyen, Eduward L., et al., eds., *Focus on Exceptional Children*. September, 1973, pp. 3–13. Reprinted in: *Alternatives for Teaching Exceptional Children*. Denver: Love Publishing Co., 1975. pp. 112–121.

Mercer, Jane R. Psychological assessment and the rights of children. In: Hobbs, Nicholas, ed. *The Classification of Children*. Vol. 1. San Francisco: Jossey-Bass, 1975.

Mercer, Jane R. Sociocultural factors in educational labeling. In: Begab, Michael J., and Richardson, Stephen A., eds. *The Mentally Retarded and Society: A Social Science Perspective*. Baltimore: University Park Press, 1975.

Mercer, Jane R., and Brown, W. C. Racial differences in IQ: Fact or artifact? In: Senna, Carl, ed. *The Fallacy of IQ*. New York: The Third Press, 1973. pp. 56–113.

Mercer, Jane R.; Butler, E. W.; and Dingman, H. F. The relationship between social-developmental performance and mental ability. *American Journal of Mental Deficiency*, 69:195–205, 1964.

★ U. S. GOVERNMENT PRINTING OFFICE : 1977 O - 245-988